

# КОНЦЕПЦИЯ СОЗДАНИЯ ИНФОРМАЦИОННО-ВЫЧИСЛИТЕЛЬНЫХ СИСТЕМ, СОДЕРЖАЩИХ ОПУБЛИКОВАННЫЕ ЧИСЛОВЫЕ МАССИВЫ ДАННЫХ

*А.И. Привезенцев<sup>1</sup>*

<sup>1</sup>Институт оптики атмосферы СО РАН, Томск  
e-mail: remake@iao.ru

## **Аннотация.**

В статье рассматривается концепция создания информационно-вычислительных web-систем в предметных областях, где существует практика опубликования рассчитанных или полученных в ходе эксперимента числовых массивов данных. Фиксирование информационных задач и создание прикладной онтологии для определенной предметной области является достаточным условием автоматизированного построения информационно-вычислительной системы заданной предметной области. Для реализации концепции создания подобных информационно-вычислительных систем описывается промежуточное программное обеспечение.

## **Введение**

В настоящее время в мире существует несколько точек зрения на способы разработки информационно-вычислительных web-систем. В данной статье будет рассматриваться концепция создания информационно-вычислительных web-систем в предметных областях, где существует практика опубликования рассчитанных или полученных в ходе физического эксперимента числовых массивов данных. Основная цель рассмотрения такой концепции состоит в новых возможностях частичной автоматизации процесса разработки и интеллектуализации информационно-вычислительных web-систем.

Для разработки web-приложений существует множество фреймворков написанных на разных языках программирования. Одним из наиболее удачных фреймворков подобного рода является Ruby on Rails[1]. В основе которого лежат следующие принципы разработки web-приложений: фреймворк предоставляет механизмы повторного использования; по умолчанию используются соглашения по конфигурации, типичные для большинства приложений. Соглашения по конфигурации позволяют автоматизировать большую часть рутинных действий и ускорить процесс разработки, тогда как явная спецификация конфигурации позволяет гибко и точно настраивать и контролировать необходимые последовательности действий в процессе разработки. Именно использование соглашений по конфигурации может частично автоматизировать процесс разработки.

Для интеллектуализации web-систем можно использовать открытые стандарты по описанию информации в сети Интернет. Наиболее популярным языком с возможностями интеллектуального логического вывода используемого для описания ресурсов является язык описания онтологий OWL[2]. Открытые программные библиотеки для работы с языком OWL наиболее полно разработаны для языка Java.

Поскольку в других языках программирования существуют свои аналоги упомянутого выше фреймворка Ruby on Rails, то наиболее предпочтительным было бы использование фреймворка на Java. Так как разработка на наиболее популярном языке программирования Java довольно консервативна и не быстра, необходимо рассмотреть альтернативы. Одним из

аналогов фреймворка Ruby on Rails, ускоряющем процесс разработки, является фреймворк Grails[3], написанный на языке Groovy, который в свою очередь основан на Java.

Таким образом, в основе разработки фреймворка для создания информационно-вычислительных web-систем можно использовать уже существующий web-фреймворк Grails, в котором уже реализована базовая функциональность web-приложения.

### **Информационная модель предметной области**

Создание онтологического описания информационно-вычислительной системы можно проводить после её создания или во время моделирования предметной области в рамках работ по её созданию, а после автоматически актуализировать его содержимое. В таком случае важным становится выбор модели предметной области, на основе которой будет создана информационно-вычислительная система.

Модель предметной области, положенная в основу систематизации данных в информационно-вычислительной системе, представляется в виде совокупности решаемых в ней задач, где каждая задача представлена в виде IPO-модели (Input-Processing-Output). Следуя этой модели предметной области, опубликованные числовые данные являются результатом решения определенной *предметной задачи*. Связка результата решения определенной предметной задачи и библиографической информации о публикации, по которой можно однозначно идентифицировать этот результат, является *информационным источником*.

На основе заданной информационной модели можно подготовить базовую прикладную онтологию, чтобы онтологическое описание предметной области происходило в терминах понятных промежуточному программному обеспечению, которое будет использоваться для создания информационно-вычислительных систем.

Информационно-вычислительная система, в основу которой положена описанная выше модель предметной области, позволяет решать типовые информационные задачи:

- создание, редактирование и удаление библиографической информации о публикуемых данных;
- программная реализация решения задачи предметной области;
- импорт и экспорт данных в информационную систему из структурированных файлов;
- редактирование и удаление данных в информационном источнике;
- проверка данных по правилам заданным предметной областью;
- расчёт дополнительных данных по правилам заданным предметной областью;
- автоматическое составление стандартных метаданных для информационных источников;
- просмотр и сравнение информационных источников в табличном и графическом видах;
- атрибутивный поиск информационных источников;
- создание составных источников информации соединением (JOIN) или объединением (UNION) данных из различных источников информации;
- декомпозиция данных источника информации по существующим в системе первичным источникам информации для проверки на ограничения существования;

– попарное вычисление среднеквадратичных отклонений в данных по заданным параметрам между сравнимыми источниками информации.

Выбор приведённых выше информационных задач и создание прикладной онтологии для определенной предметной области будет являться достаточным условием частичного автоматизированного построения информационно-вычислительной системы заданной предметной области.

Для реализации концепции создания подобных информационно-вычислительных систем разрабатывается промежуточное программное обеспечение SPIN-фреймворк.

Созданные онтологии могут использоваться для машинной систематизации и интерпретации знаний, для интеграции знаний в другие смежные предметные области, а также организации семантического поиска. Классы базовой прикладной онтологии SPIN показаны на рис. 1.



Рис. 1. Классы базовой прикладной онтологии SPIN

Свойства базовой прикладной онтологии SPIN показаны на рис. 2.



Рис. 2. Свойства базовой прикладной онтологии SPIN

В качестве примера составления прикладной онтологии предметной области взята молекулярная спектроскопия[4]. Для молекулярной спектроскопии характерно большое количество числовых массивов спектральных данных, имеющих большие объёмы. Информационная модель предметной области выделяет следующие подклассы научных задач (ScientificProblem). Прямые задачи молекулярной спектроскопии связаны с расчетами из первых принципов фундаментальных характеристик молекул, таких как уровни энергии молекул, частоты перехода, коэффициенты Эйнштейна и т.д. Обратные задачи молекулярной спектроскопии связаны с обработкой данных измерений спектральных функций, что позволяет в дальнейшем при машинной обработке классифицировать их выходные данные как экспериментальные.

К классам элементарных прямых задач, используемых нами для проектирования информационной системы, относятся следующие классы задач: задача определения физических характеристик изолированной молекулы (T1); задача определения параметров спектральной линии изолированной молекулы (T2); задача определения параметров контура спектральной линии (T3); задача расчета спектральных функций (T4).

К классам элементарных обратных задач, используемых нами для проектирования информационной системы, относятся следующие классы: задача измерения спектральных функций (E1); задача приписывания квантовых чисел спектральным линиям (T5); задача определения коэффициентов Эйнштейна (T6); задача определения уровней энергии изолированной молекулы (T7).

На рис . 3 показаны классы прикладной онтологии по молекулярной спектроскопии

для задачи T1.

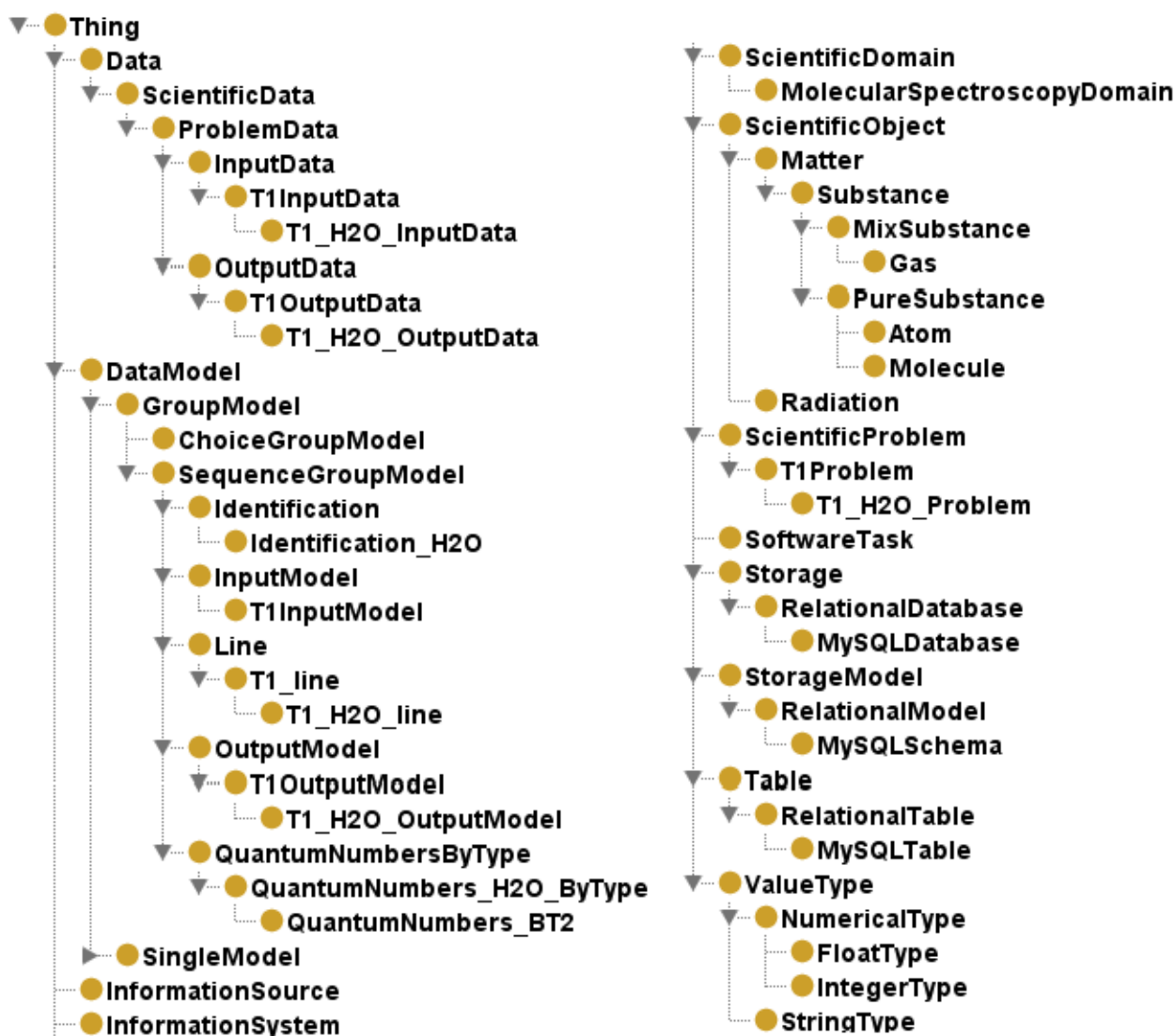


Рис. 3. Классы прикладной онтологии по молекулярной спектроскопии для задачи T1

Для создания прикладных онтологий для информационно-вычислительных систем разрабатывается графический web-интерфейс, облегчающий процесс разработки на базе онтологии SPIN.

### Диаграммы прецедентов информационно-вычислительной системы

Одним из способов моделирования поведения в UML[5] используются диаграммы прецедентов использования. На рис. 4 представлены диаграмма прецедентов использования SPIN-фреймворка. Моделированием предметной области на языке OWL занимается эксперт, после создания прикладной онтологии предметной области автоматически создаются реляционная база данных и XML-схемы для работы с числовыми массивами данных. Если базовая функциональность SPIN-фреймворка для предметной области не устраивает эксперта, то разработчик доводит её до нужного состояния.

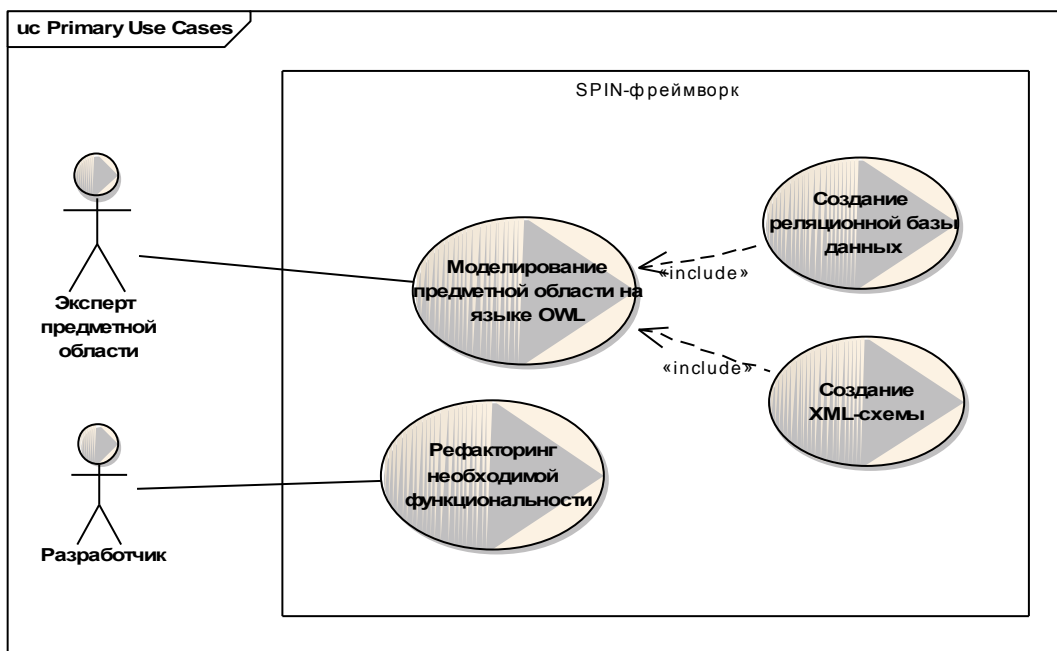


Рис. 4. Диаграмма прецедентов использования SPIN-фреймворка

На рис. 5 представлены диаграмма прецедентов использования информационной системы по молекулярной спектроскопии с точки зрения пользователя и решаемых им информационных задач.

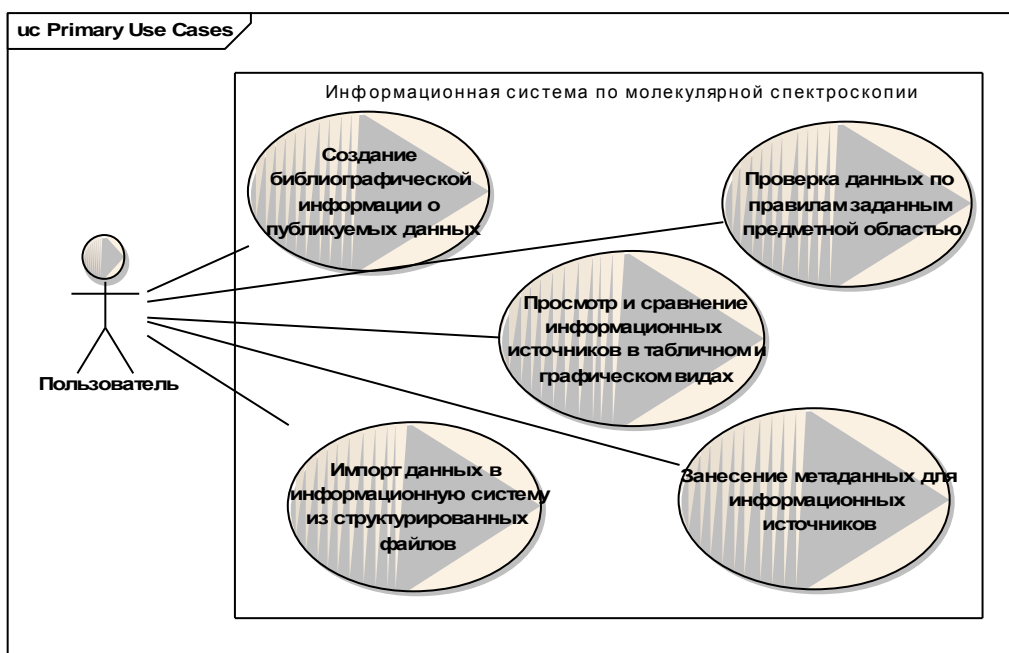


Рис. 5. Диаграмма прецедентов использования информационной системы по молекулярной спектроскопии

## Заключение

Основная цель, рассматриваемой концепции создания информационно-вычислительных web-систем в предметных областях, где существует практика опубликования рассчитанных или полученных в ходе физического эксперимента числовых

массивов данных, может быть достигнута с помощью SPIN-фреймворка на базе web-фреймворка Grails и прикладных онтологий предметных областей на базе онтологии SPIN. Основным преимуществом использования онтологии на языке OWL DL для моделирования предметной области является возможность использования универсального формата обмена знаниями в Web с явной спецификацией знаний. Кроме этого в прикладной онтологии возможна машинная логическая проверка с помощью машины вывода на достоверность и неполноту информации о моделируемой информационно-вычислительной системе.

## Литература

- [1]. Ruby on Rails / David Heinemeier Hansson. – Электрон. текстовые дан. – 2003. – Режим доступа: <http://rubyonrails.org/>.
- [2]. OWL Web Ontology Language Semantics and Abstract Syntax / edited by P. F. Patel-Schneider, I. Horrocks. – Электрон. текстовые дан. – 2004. – Режим доступа: <http://www.w3.org/TR/owl-semantics/>. – W3C Recommendation.
- [3]. Grails / SpringSource. – Электрон. текстовые дан. – 2009. – Режим доступа: <http://www.grails.org/>.
- [4]. Козодоев, А. В. Информационная система для решения задач молекулярной спектроскопии. 3. Уровни энергии молекул / А. В. Козодоев, А. И. Привезенцев, А. З. Фазлиев // Оптика атмосферы и океана. – 2007. – Т. 20. – № 9. – С. 805-809.
- [5]. Мацяшек, Л. Анализ и проектирование информационных систем с помощью UML 2.0 / Л. Мацяшек. – М.: Вильямс, 2008. – 816 с.